Lecture15

Monday, November 8, 2021 4:28 PM

Lecture15

Linear Algebra

Samira Hossein Ghorban
s.hosseinghorban@ipm.ir

Fall, 2021

System Equations

We want to fit a plane $y = C + Dt + Ez$ to the four points

$$y = 3 \quad \text{at } (t, z) = (1, 1), \qquad y = 6 \quad \text{at } (t, z) = (0, 3),$$

$$y = 5 \quad \text{at } (t, z) = (2, 1), \qquad y = 0 \quad \text{at } (t, z) = (0, 0),$$

- $y = C + Dt + Ez$

$D = 1 \Leftarrow D + 2 = 3$

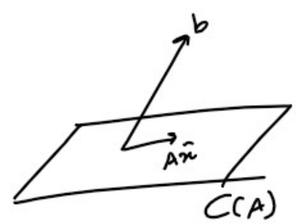
$D = 3/2 \Leftarrow 2D + 2 = 5$

$$\Leftarrow \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \\ 1 & 0 & 0 \end{bmatrix}}_{A} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \overset{b}{\begin{bmatrix} 3 \\ 5 \\ 6 \\ 0 \end{bmatrix}}$$

$3 = C + D + E$
$5 = C + 2D + E$
$6 = C + 0D + 3E$
$0 = C + 0D + 0E$

- $C = 0$, $D = ?$, $E = 2$.

$\boxed{C = 0}$
$3E = 6$ $\boxed{E = 2}$



$\|b - A\hat{x}\|^2$

$(1, 1) \rightarrow y = 3$

$(0, 3, 6)$
t z y

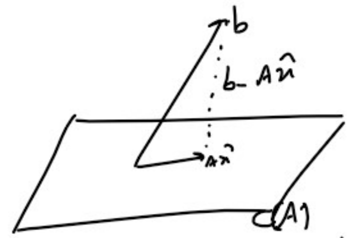
Least Square

- Find C and D such that the following data are fitted by $b = C + tD$.

t	b	t	b
1.0	−1.945	3.2	0.764
1.2	−1.253	3.4	0.532
1.4	−1.140	3.6	1.073
1.6	−1.087	3.8	1.286
1.8	−0.760	4.0	1.502
2.0	−0.682	4.2	1.582
2.2	−0.424	4.4	1.993
2.4	−0.012	4.6	2.473
2.6	−0.190	4.8	2.503
2.8	0.452	5.0	2.322
3.0	0.337		

Least squares problems with several variables

- So, instead of one column and one unknown x, the matrix now has n columns. The number m of observations is still larger than the number n of unknowns.
- So it must be expected that $Ax = b$ will be inconsistent. Probably, there will not exist a choice of x that perfectly fits the data b.
- In other words, the vector b probably will not be a combination of the columns of A; it will be outside the column space.
- Solution: finding $\hat{x}$ which minimizes $E = |b - Ax|$.
- Equivalently, finding $\hat{x}$ such that the error vector $e = b - A\hat{x}$ be perpendicular to that space $C(A)$.



$b - A\hat{x} \in (C(A))^{\perp} = N(A^T)$

Least squares problems with several variables

- Finding $\hat{x}$ such that the error vector $e = b - A\hat{x}$ be perpendicular to that space $C(A)$ is equivalent to

$$A^T(b - A\hat{x}) = 0$$
$$A^T A\hat{x} = A^T b$$

- If $A^T A$ is invertible, then the best square estimation :

$$\hat{x} = (A^T A)^{-1} A^T b.$$

Weighted Least Squares

- Now suppose that the two observations are not trusted to the same degree.
- A simple least-squares problem is to estimate two observations $x = b_1$ and $x = b_2$.
- if $b_1 \neq b_2$, then we are faced with an inconsistent system of two equations in one unknown:

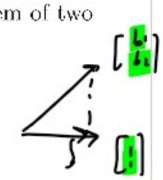
$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

- Thus

$$E^2 = (x - b_1)^2 + (x - b_2)^2$$

- The best square estimation:

$$\hat{x} = \frac{b_1 + b_2}{2}.$$



$P = s\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\dfrac{b_1 + b_2}{2}$

Weighted Least Squares

- Now, suppose that the the value $x = b_1$ may be obtained from a more accurate source from a larger $x = b_2$.
- The simplest compromise is to attach different weights w_1 and w_2 and choose the $\hat{x}_W$ that minimizes the weighted sum of squares:

$$E^2 = w_1^2 (x - b_1)^2 + w_2^2 (x - b_2)^2.$$

- The best square estimation :

$$\hat{x} = \frac{w_1^2 b_1 + w_2^2 b_2}{w_1^2 + w_2^2}.$$

- Generally:

$$(A^T W^T W A)\hat{x}_W = A^T W^T W b.$$

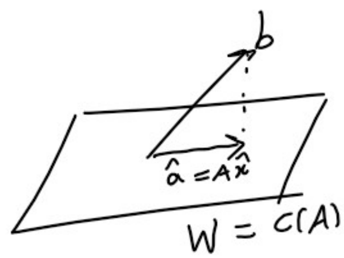
$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} x = \begin{bmatrix} w_1 b_1 \\ w_2 b_2 \end{bmatrix}$

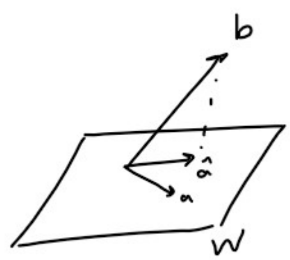
$\hat{x} = \dfrac{w_1^2 b_1 + w_2^2 b_2}{w_1^2 + w_2^2}$

$Ax = b$

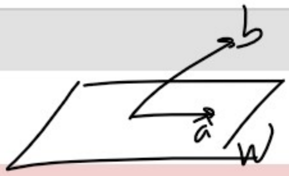
$b \notin C(A)$

$\|b - \hat{a}\| \leq \|b - a\|$





Approximation



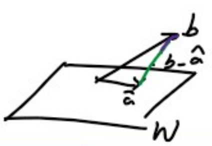
A best approximation

A **best approximation** to b by vectors in W is a vector $a in W$ such that

$$\|b - \hat{a}\| \leqslant \|b - a\|$$

for each vector $a \in W$. The vector $\hat{a} \in W$ is called a best approximation to b in W.

$\|b - \hat{a}\|$



$\|b - \hat{a}\| \leq \|b - a\|$

برهان قضیه: (ا) $(\Leftarrow)$ فرض کنید $\hat{a}$ در W نزدیک‌ترین تقریب توسط بردارهای W برای b باشد. به ازای هر $c \in W$

$\|b - c\|^2 = \|b - \hat{a} + \hat{a} - c\|^2$

$= \|b - \hat{a}\|^2 + \|\hat{a} - c\|^2 + 2\langle b - \hat{a}, \hat{a} - c \rangle$

Approximation

Approximation

Theorem

Let W be a subspace of an inner product space V and let b be a vector in V.

(i) *The vector $\hat{a}$ in W is a best approximation to b by vectors in W if and only if $b - \hat{a}$ is orthogonal to every vector in W.*

(ii) *If a best approximation to b by vectors in W exists, it is unique.*

(iii) *If W is finite-dimensional and $\{v_1, \ldots, v_n\}$ is any orthonormal basis for W, then the vector*

$$\hat{a} = \sum_{k=1}^{n} \langle b, v_k \rangle v_k$$

is the (unique) best approximation to b by vectors in W.

Inner product on function spaces

- Let V be a space of all continuous functions on the interval $[0, 2\pi]$.
- How can we measure their "length" in the space?
- The same idea of replacing summation by integration produces the inner product of two functions.
- Let V be a space of all continuous functions on the interval $[0, 2\pi]$ and $f, g \in V$, and define

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx.$$

- thus,

$$\|f\|^2 = \int_0^{2\pi} f(x)^2 dx.$$

Inner product on function spaces

- Consider $\sin(x), \cos(x) \in V$.

$$\|\sin(x)\|^2 = \int_0^{2\pi} \sin^2(x)dx = \frac{1}{2}\left(2\pi - \int_0^{2\pi} \cos 2x\right) = \pi.$$

$$\|\cos(x)\|^2 = \int_0^{2\pi} \cos^2(x)dx = \frac{1}{2}\left(2\pi + \int_0^{2\pi} \cos 2x\right) = \pi.$$

$$\langle \sin(x), \cos(x) \rangle = \int_0^{2\pi} \sin x \cos x dx = \frac{1}{2}\int_0^{2\pi} \sin 2x dx = 0$$

- Thus, the functions $\sin(x)$ and $\cos(x)$ are orthogonal. Thus, $\frac{1}{\sqrt{\pi}}\sin(x)$ and $\frac{1}{\sqrt{\pi}}\cos(x)$ are orthonormal.
- so, $\frac{m}{\sqrt{\pi}} \sin mx$ and $\frac{n}{\sqrt{\pi}} \cos nx$ are orthonormal.

$\{1, x, x^2, x^3, x^4, x^5\}$

Gram-Schmidt for Functions

$= \|b - \hat{a}\|^2 + \|\hat{a} - c\|^2 + 2\langle b - \hat{a}, \hat{a} - c\rangle$

$\|b - \hat{a}\| \leq \|b - c\|$

$\Rightarrow \|\hat{a} - c\|^2 + 2\langle b - \hat{a}, \hat{a} - c\rangle \geq 0 \quad \forall c \in W$

$t = \hat{a} - c$

$\forall t \in W \quad \|t\|^2 + 2\langle b - \hat{a}, t\rangle \geq 0$

$t = \frac{\langle b - \hat{a}, \hat{a} - c\rangle}{\|\hat{a} - c\|^2}(\hat{a} - c)$

$\|t\|^2 + 2\langle b - \hat{a}, t\rangle =$

$\frac{\langle b - \hat{a}, \hat{a} - c\rangle^2}{\|\hat{a} - c\|^2} - 2\frac{\langle b - \hat{a}, \hat{a} - c\rangle^2}{\|\hat{a} - c\|^2}$

$= -\frac{\langle b - \hat{a}, \hat{a} - c\rangle^2}{\|\hat{a} - c\|^2} \geq 0$

$\Downarrow$

$\langle b - \hat{a}, \hat{a} - c\rangle^2 = 0$

$\Downarrow$

$\langle b - \hat{a}, \hat{a} - c\rangle = 0$

$\Downarrow \quad \forall c \in W$

$b - \hat{a} \in W^\perp$

(⇒) فرض کنید $b - \hat{a}$ بر هر بردار W عمود است $(b - \hat{a} \in W^\perp)$

آنگاه $\hat{a}$ بهترین تقریب برای b است ≡

$\|b - \hat{a}\| \leq \|b - c\| \quad \forall c \in W$

$\hat{a} \in W$

$b - c = (b - \hat{a}) + (\hat{a} - c)$

$\|b - c\|^2 = \|b - \hat{a}\|^2 + \|\hat{a} - c\|^2 + 2\langle b - \hat{a}, \underbrace{\hat{a} - c}_{W}\rangle$

$= \|b - \hat{a}\|^2 + \|\hat{a} - c\|^2$

ب) بهترین تقریب یکتاست:

فرض کنید $\hat{a}$ و $\hat{\hat{a}}$ دو بهترین تقریب برای بردار b توسط W باشند.

پس بنا بر الف $b - \hat{a} \in W^\perp$ و $b - \hat{\hat{a}}$

$\Downarrow$

$(b - \hat{\hat{a}}) - (b - \hat{a}) \in W^\perp$

$\hat{a} - \hat{\hat{a}} \in W^\perp$

$\{1, x, x^2, x^3, x^4, x^5\}$

Gram-Schmidt for Functions

v_1, v_2, v_3

- Are $\{1, x, x^2\}$ are orthogonal?
- we have
$$\langle 1, x\rangle = \int_{-1}^{1} x dx = 0.$$

$$\langle 1, x^2\rangle = \int_{-1}^{1} x^2 dx = \frac{x^3}{3}\Big|_{-1}^{1} = \frac{1}{3}(1-(-1))$$

$\frac{2}{3} \neq 0$

- Let $q_1 = \frac{1}{\sqrt{\langle 1,1\rangle}}$ and $q_2 = \frac{x}{\sqrt{\langle x,x\rangle}}$.

$\{1, x, x^2 - 1/3\}$

$$Q_3 = x^2 - \frac{\langle x^2, 1\rangle}{\langle 1, 1\rangle}1 - \frac{\langle x^2, x\rangle}{\langle x, x\rangle}x = x^2 - \frac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} x dx} - 0 = x^2 - \frac{1}{3}$$

$$Q_3 = v_3 - \langle v_3, q_1\rangle q_1 - \langle v_3, q_2\rangle q_2$$

- The polynomials $1, x, x^2 - \frac{1}{3}$ are orthogonal to each other over the interval $-1 \leqslant x \leqslant 1$ they are called ***Legendre polynomials***.

$$Q_3 = x^2 - \left\langle x^2, \frac{1}{\sqrt{\langle 1,1\rangle}}\right\rangle \frac{1}{\sqrt{\langle 1,1\rangle}} - \left\langle x^2, \frac{x}{\sqrt{\langle x,x\rangle}}\right\rangle \frac{x}{\sqrt{\langle x,x\rangle}}$$

Best Straight Line

- Suppose we want to approximate $y = x^5$ by a straight line $\underline{C + Dx}$ between $x = 0$ and $x = 1$.

W متناظر به نزدیکترین ها است

- There are at least three ways of finding that line.
  1. Solve $\begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = x^5$ by least squares.
$$E^2 = \| x^5 - (C + Dx) \|^2$$
  2. Minimize
$$E^2 = \int_0^1 (x^5 - C - Dx)^2 dx.$$
  3. Apply Gram-Schmidt!

...

$v =$

$[\ ]$ $\int f(x) g(x)$

Thank You!

$$\underline{(b - \hat{\hat{a}}) - (b - \hat{a})} \in W^{\perp}$$

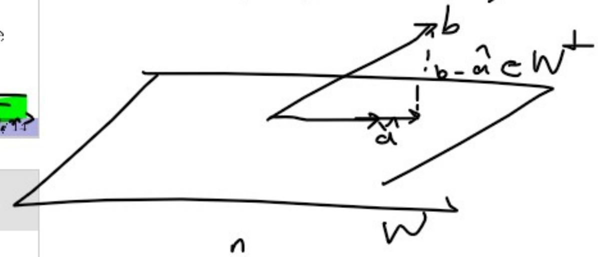
$$\hat{\hat{a}} - \hat{a} \in W^{\perp}$$

$$\hat{\hat{a}} - \hat{a} \in W \Leftarrow \hat{a}, \hat{\hat{a}} \in W$$

$$W \cap W^{\perp} = \{0\}$$

$$\Rightarrow \hat{\hat{a}} - \hat{a} = 0 \Rightarrow \underline{\underline{\hat{\hat{a}} = \hat{a}}}$$

$$W = \mathrm{span}(\{v_1, \ldots, v_n\})$$ فرض کنیم (ح.



$$\hat{a} = \sum_{i=1}^{n} d_i v_i$$

â را باید بیابیم که $\underline{b - \hat{a} \in W^{\perp}}$ یعنی $b - \hat{a}$ بر هر بردار W عمود باشد

$\Downarrow$

$$\forall 1 \leq k \leq n \quad \langle b - \hat{a}, v_k\rangle = 0$$

$$0 = \langle b - \hat{a}, v_k\rangle$$
$$= \langle b, v_k\rangle - \langle \hat{a}, v_k\rangle$$
$$= \langle b, v_k\rangle - \left\langle \sum_{i=1}^{n} d_i v_i, v_k\right\rangle$$
$$= \langle b, v_k\rangle - \sum_{i=1}^{n} d_i \langle v_i, v_k\rangle$$

$i \neq k$
$0 = \langle v_i, v_k\rangle$
$1 = \langle v_k, v_k\rangle$

$$= \langle b, v_k\rangle - d_k$$

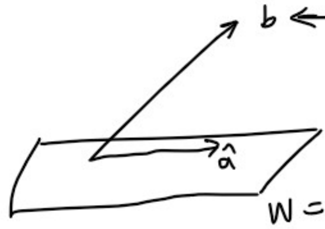
$$\Rightarrow d_k = \langle b, v_k\rangle$$

$$\Rightarrow \hat{a} = \sum_{i=1}^{n} \langle b, v_j\rangle v_i$$

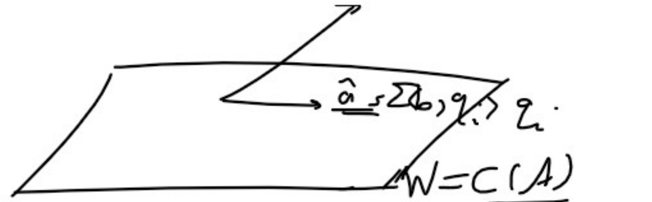
$\nearrow b$

تمرین: تقریب $y=x^5$ در بازه $0\le x\le 1$ با خط راست $\underline{\underline{y=Dx+C}}$ را بیابید.

$b \leftarrow x^5$



$W=\operatorname{span}\left(\{1,x\}\right)$



$q_1, q_2 \cdots q_r$

$Ax=b$

A^TA

$C(A)$

$$\langle 1,x\rangle=\int_0^1 x\,dx=\frac{x^2}{2}\Big|_0^1=1/2\neq 0$$

$\{1,x\}$ که متعامد برای W نیست.

از روند گرام-اشمیت استفاده می‌کنیم:

$$\langle 1,1\rangle=\int_0^1 dx=x\Big|_0^1=1$$

$$q_1=\frac{1}{\sqrt{\langle 1,1\rangle}}=1$$

$\{q_1, \to q_n\}$

$$\hat{a}=\sum_{i=1}^{n}\langle \underline{\underline{b}},q_i\rangle q_i$$

$$Q_2=x-\langle x,q_1\rangle q_1=x-\langle x,1\rangle 1$$

$$=x-\int_0^1 x\,dx=x-\frac{x^2}{2}\Big|_0^1=x-1/2$$

$$\|Q_2\|^2=\langle x-1/2,\,x-1/2\rangle=\int_0^1(x-1/2)^2dx=\int_0^1\left(x^2-x+1/4\right)dx=\frac{x^3}{3}-\frac{x^2}{2}+1/4\,x\Big|_0^1$$

$$\overset{4-6+3}{=1/3-1/2+1/4}=\frac{1}{12}$$

$$q_2=\frac{1}{\sqrt{12}}(x-1/2)$$

$$\hat{C}+\hat{D}x=\langle x^5,1\rangle 1+\frac{\overbrace{\langle x^5,\,x-1/2\rangle}^{\int_0^1 (x^6-1/2\,x^5)dx}}{\sqrt{12}}\,\frac{(x-1/2)}{\sqrt{12}}$$

$$=\underbrace{\int_0^1 x^5\,dx}_{1/6}+\frac{1}{12}\Big(\underbrace{\underbrace{x^7\Big|_0^1}-1/6\,x^6\Big|_0^1}_{1/7-1/36}\Big)(x-1/2)$$

$$=1/6+\frac{5}{7}(x-1/2)$$

پایه متعامد $\underline{\underline{W}}$

$$E(C,D)=\|x^5-C-Dx\|^2=\int_0^1\left(x^5-C-Dx\right)^2dx$$

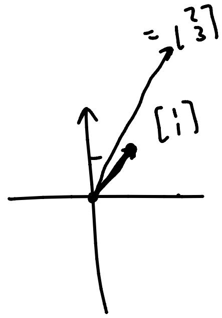
$$=\int_0^1\left(x^{10}+C^2+D^2x-2Cx^5-2Dx^6+2CDx\right)dx$$

$$= \frac{1}{7} - \frac{2}{4}C - \frac{2}{5}D + C^2 + CD + \frac{1}{3}D^2$$

$$\frac{\partial E}{\partial C} = -\frac{2}{4} + 2C + D = 0 \implies C + \frac{1}{2}D = \frac{1}{4}$$

$$\frac{\partial E}{\partial D} = -\frac{2}{5} + C + \frac{2}{3}D = 0 \implies \frac{1}{2}C + \frac{1}{3}D = \frac{1}{5}$$

$$\implies \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/5 \end{bmatrix}$$



$y =$